

DATA-ADAPTIVE REDUCED-DIMENSION ROBUST BEAMFORMING ALGORITHMS

Samuel D. Somasundaram*, Nigel H. Parsons

General Sonar Studies
Thales Underwater Systems
Cheadle Heath, Stockport, Cheshire, U.K.

Peng Li, Rodrigo C. de Lamare

Communications Research Group
Department of Electronics
University of York, U.K.

ABSTRACT

We present low complexity, quickly converging robust adaptive beamformers that combine robust Capon beamformer (RCB) methods and data-adaptive Krylov subspace dimensionality reduction techniques. We extend a recently proposed reduced-dimension RCB framework, which ensures proper combination of RCBs with any form of dimensionality reduction that can be expressed using a full-rank dimension reducing transform, providing new results for data-adaptive dimensionality reduction. We consider Krylov subspace methods computed with the Powers-of-R (PoR) and Conjugate Gradient (CG) techniques, illustrating how a fast CG-based algorithm can be formed by beneficially exploiting that the CG-algorithm diagonalizes the reduced-dimension covariance. Our simulations show the benefits of the proposed approaches.

Index Terms— Robust adaptive beamforming, dimensionality reduction, Krylov subspace methods.

1. INTRODUCTION AND PRELIMINARIES

When implementing adaptive beamforming on arrays with large aperture and many elements that operate in dynamic environments, reduced-dimension techniques are often needed to speed-up the convergence of beamforming algorithms and reduce the computational complexity [1]. This is of fundamental importance in applications found in passive sonar and radar systems. Furthermore, robust adaptive techniques are often required to alleviate the deleterious effects of array steering vector (ASV) mismatch, e.g., caused by calibration and pointing errors. A popular class of these are the robust Capon beamformers (RCBs) that exploit ellipsoidal, including spherical, uncertainty sets of the ASV [2–6]. In [1, 7], a framework for combining reduced-dimension and RCB techniques was derived, allowing rapidly converging, low complexity robust adaptive reduced-dimension robust Capon beamformers (RDRCBs) to be formed. A key contribution of that work was the derivation of a complex propagation theorem that allows a reduced-dimension ellipsoid to be derived from an element-space ellipsoid and any full-rank dimension reducing transform (DRT). The reduced-dimension ellipsoid may then be exploited by using an RCB in the reduced-dimension space. In [1, 7], only data-independent dimensionality reduction was considered. Here, we extend the framework developed in [1, 7] to data-adaptive dimensionality reduction, providing new results useful for exploiting a variety of scenarios that occur in practical applications of robust beamforming algorithms.

The problem under consideration is the design of RDRCBs that are suitable for large arrays. We consider Krylov subspace

techniques [8], [9], [10], [11], [12] for data-adaptive dimensionality reduction which are computed by the Powers-of-R (PoR) [8], [9], [10], [13], [14] and Conjugate-Gradient (CG) [12], FLW10 algorithms. We then develop RCB versions of the PoR and CG algorithms for large arrays. We present a CG-based technique can exploit the fact that it results in a diagonal reduced-dimension sample covariance matrix to give particularly low-complexity data-adaptive beamforming algorithms. Scenarios with large planar arrays are investigated along with both non-degenerate ellipsoidal uncertainty and spherical uncertainty sets.

In the following, $E\{\cdot\}$, $(\cdot)^T$, $(\cdot)^H$, $(\cdot)^{-1}$ and $(\cdot)^\dagger$ denote the expectation, transpose, Hermitian transpose, inverse and Moore-Penrose pseudo-inverse operators, respectively. Furthermore, $\|\cdot\|_2$, $\mathbf{N}_\mathbf{X}^\perp$, $\Pi_\mathbf{X}$ and $\Pi_\mathbf{X}^\perp$ denote the two-norm, a basis for the left null-space of \mathbf{X} , the orthogonal projector onto the range space of \mathbf{X} and the orthogonal projector onto the space perpendicular to the range space of \mathbf{X} , respectively. Moreover, $\mathbf{X} \geq 0$ or $\mathbf{X} > 0$ mean that the Hermitian matrix \mathbf{X} is +ve semi-definite or +ve definite.

1.1. Robust Capon Beamforming

We model the k th element-space array snapshot $\mathbf{x}_k \in \mathbb{C}^M$ as

$$\mathbf{x}_k = \mathbf{a}_0 s_{0,k} + \mathbf{n}_k, \quad (1)$$

where \mathbf{a}_0 , $s_{0,k}$ and \mathbf{n}_k denote the true signal-of-interest (SOI) ASV, the SOI complex amplitude and an additive zero-mean complex Gaussian vector that incorporates the noise and the interference. Assuming $s_{0,k}$ is zero mean and uncorrelated with \mathbf{n}_k , the array covariance matrix can be written as $\mathbf{R}_\mathbf{x} = E\{\mathbf{x}_k \mathbf{x}_k^H\} = \sigma_0^2 \mathbf{a}_0 \mathbf{a}_0^H + \mathbf{Q}_\mathbf{x}$, where $\mathbf{R}_\mathbf{x} > 0$, $\sigma_0^2 = E\{|s_{0,k}|^2\}$ is the SOI power and $\mathbf{Q}_\mathbf{x} = E\{\mathbf{n}_k \mathbf{n}_k^H\}$ is the noise plus interference covariance matrix. In practice, $\mathbf{R}_\mathbf{x}$ is often replaced by the sample covariance matrix (SCM)

$$\hat{\mathbf{R}}_\mathbf{x} = \frac{1}{K} \sum_{k=1}^K \mathbf{x}_k \mathbf{x}_k^H, \quad (2)$$

formed from K snapshots. In [3] (see, also [6]), RCBs were derived by solving $\max_{\sigma^2, \mathbf{a}} \sigma^2$ s.t. $\mathbf{R}_\mathbf{x} - \sigma^2 \mathbf{a} \mathbf{a}^H \geq 0$, $\mathbf{a} \in \mathcal{E}_M(\bar{\mathbf{a}}, \mathbf{E})$, which can be reduced to [3]

$$\min_{\mathbf{a}} \mathbf{a}^H \mathbf{R}_\mathbf{x}^{-1} \mathbf{a} \quad \text{s.t.} \quad \mathbf{a} \in \mathcal{E}_M(\bar{\mathbf{a}}, \mathbf{E}). \quad (3)$$

The M -dimensional element-space ellipsoid $\mathcal{E}_M(\bar{\mathbf{a}}, \mathbf{E})$ is parameterized by $\bar{\mathbf{a}}$, which often represents the assumed ASV, and $\mathbf{E} \geq 0 \in \mathbb{C}^{M \times M}$, and can be written as

$$\mathcal{E}_M(\bar{\mathbf{a}}, \mathbf{E}) = \left\{ \mathbf{a} \in \mathbb{C}^M \mid [\mathbf{a} - \bar{\mathbf{a}}]^H \mathbf{E} [\mathbf{a} - \bar{\mathbf{a}}] \leq 1 \right\}. \quad (4)$$

*This work was supported by MOD/DTIC under contract RT/COM/4/032.

For non-degenerate sets, $\mathbf{E} > 0$. To solve (3), we assume that

$$\bar{\mathbf{a}}^H \mathbf{E} \bar{\mathbf{a}} > 1 \quad (5)$$

When $\mathbf{E} = (1/\epsilon)\mathbf{I}$, (4) reduces to a spherical uncertainty set, $\|\mathbf{a} - \bar{\mathbf{a}}\|_2^2 \leq \epsilon$, with radius $\sqrt{\epsilon}$ and (5) becomes $\|\bar{\mathbf{a}}\|_2^2 > \epsilon$. For non-degenerate ellipsoids, we can factor $\mathbf{E} = \mathbf{E}^{\frac{H}{2}} \mathbf{E}^{\frac{1}{2}}$ and form $\check{\mathbf{a}} = \mathbf{E}^{\frac{1}{2}} \bar{\mathbf{a}}$, $\check{\mathbf{a}} = \mathbf{E}^{\frac{1}{2}} \bar{\mathbf{a}}$ and $\check{\mathbf{R}} = \mathbf{E}^{\frac{1}{2}} \mathbf{R} \mathbf{E}^{\frac{H}{2}}$. Then, (3) can be re-written using the following spherical constraint [4]

$$\min_{\check{\mathbf{a}}} \check{\mathbf{a}}^H \check{\mathbf{R}}^{-1} \check{\mathbf{a}} \text{ s.t. } \|\check{\mathbf{a}} - \check{\mathbf{a}}\|_2^2 \leq 1. \quad (6)$$

As shown in [4], (6) can be solved via the eigenvalue decomposition (EVD) of $\check{\mathbf{R}}$, where computing the EVD is the most computationally expensive step. Denoting $\hat{\check{\mathbf{a}}}$ as the solution to (6), the solution to (3) is formed as $\hat{\mathbf{a}}_{0,\text{RCB}} = \mathbf{E}^{-\frac{1}{2}} \hat{\check{\mathbf{a}}}$. The RCB power estimate is formed as $\hat{\sigma}_{0,\text{RCB}}^2 = \frac{\|\hat{\mathbf{a}}_{0,\text{RCB}}\|_2^2/M}{\hat{\mathbf{a}}_{0,\text{RCB}}^H \mathbf{R}_x^{-1} \hat{\mathbf{a}}_{0,\text{RCB}}}$ and the weight vector as $\hat{\mathbf{w}}_{\text{RCB}} = \frac{\mathbf{R}_x^{-1} \hat{\mathbf{a}}_{0,\text{RCB}}}{\hat{\mathbf{a}}_{0,\text{RCB}}^H \mathbf{R}_x^{-1} \hat{\mathbf{a}}_{0,\text{RCB}}}$.

2. ROBUST CAPON BEAMFORMING FRAMEWORK WITH DATA-ADAPTIVE REDUCED-DIMENSION

In reduced-dimension methods, the k th element-space snapshot, $\mathbf{x}_k \in \mathbb{C}^M$, is projected onto an N -dimensional subspace (with $N < M$) using a DRT $\mathbf{D} \in \mathbb{C}^{M \times N}$, yielding the reduced-dimension snapshot, $\mathbf{y}_k = \mathbf{D}^H \mathbf{x}_k$, where $\mathbf{y}_k \in \mathbb{C}^N$. As shown in [1, 7], this leads to the following RDRCB problem $\max_{\sigma^2, \mathbf{b}} \sigma^2$ s.t. $\mathbf{R}_y - \sigma^2 \mathbf{b} \mathbf{b}^H \geq 0$, $\mathbf{b} \in \mathcal{E}_N(\bar{\mathbf{b}}, \mathbf{F})$, where $\mathbf{b} = \mathbf{D}^H \mathbf{a}$, $\mathbf{R}_y = \mathbf{D}^H \mathbf{R}_x \mathbf{D}$ and $\mathcal{E}_N(\bar{\mathbf{b}}, \mathbf{F})$ denote the reduced-dimension ASV, covariance and uncertainty ellipsoid, respectively, which can be reduced to

$$\min_{\mathbf{b}} \mathbf{b}^H \mathbf{R}_y^{-1} \mathbf{b} \text{ s.t. } \mathbf{b} \in \mathcal{E}_N(\bar{\mathbf{b}}, \mathbf{F}). \quad (7)$$

The following theorem is used to derive $\mathcal{E}_N(\bar{\mathbf{b}}, \mathbf{F})$.

Propagation Theorem: [1, 7] The propagation of the element-space ellipsoid (4), with $\mathbf{E} \geq 0 \in \mathbb{C}^{M \times M}$, through the mapping $\mathbf{D}^H \mathbf{a} - \mathbf{I}_N \mathbf{b} = \mathbf{0}$, where $\mathbf{D} \in \mathbb{C}^{M \times N}$ has full column rank, yields the ellipsoid $\mathcal{E}_N(\bar{\mathbf{b}}, \mathbf{F})$ [see (4)] with

$$\begin{aligned} \bar{\mathbf{b}} &= \mathbf{D}^H \bar{\mathbf{a}} \\ \mathbf{F} &= \mathbf{D}^\dagger (\mathbf{E} - \mathbf{E} \mathbf{N}_D^L [(\mathbf{N}_D^L)^H \mathbf{E} \mathbf{N}_D^L]^\dagger (\mathbf{N}_D^L)^H \mathbf{E}) (\mathbf{D}^\dagger)^H. \end{aligned} \quad (8)$$

For data-adaptive dimensionality reduction, $\bar{\mathbf{b}}$ and \mathbf{F} need updating each time the DRT is updated. If we use (9) for updating, then we observe that \mathbf{N}_D^L , $[(\mathbf{N}_D^L)^H \mathbf{E} \mathbf{N}_D^L]^\dagger$ and \mathbf{D}^\dagger need calculating, which are expensive operations. Fortunately, if the original element-space ellipsoid is non-degenerate, such that $\mathbf{E} > 0$, we can simplify (9). Then, $[(\mathbf{N}_D^L)^H \mathbf{E} \mathbf{N}_D^L]^\dagger = [(\mathbf{N}_D^L)^H \mathbf{E} \mathbf{N}_D^L]^{-1}$ and we can write

$$\begin{aligned} \mathbf{F} &= \mathbf{D}^\dagger \mathbf{E}^{\frac{1}{2}} \Pi_{\mathbf{E}^{\frac{1}{2}} \mathbf{N}_D^L}^\perp \mathbf{E}^{\frac{1}{2}} (\mathbf{D}^\dagger)^H \\ &= \mathbf{D}^\dagger \mathbf{E}^{\frac{1}{2}} \Pi_{\mathbf{E}^{-\frac{1}{2}} \mathbf{D}}^\perp \mathbf{E}^{\frac{1}{2}} (\mathbf{D}^\dagger)^H \\ &= [\mathbf{D}^H \mathbf{E}^{-1} \mathbf{D}]^{-1}, \end{aligned} \quad (10)$$

where $\Pi_{\mathbf{E}^{\frac{1}{2}} \mathbf{N}_D^L}^\perp = \mathbf{I} - \mathbf{E}^{\frac{1}{2}} \mathbf{N}_D^L [(\mathbf{N}_D^L)^H \mathbf{E} \mathbf{N}_D^L]^{-1} (\mathbf{N}_D^L)^H \mathbf{E}^{\frac{1}{2}}$. The inverse \mathbf{E}^{-1} can be computed offline and therefore, the online computation of \mathbf{F} reduces to the computation of an $N \times N$ inverse. Note that, in general, we will need to compute $\mathbf{F}^{\frac{1}{2}}$, $\mathbf{F}^{\frac{H}{2}}$ and

$\mathbf{F}^{-\frac{1}{2}}$ [see Section 1.1], which can all be obtained from the EVD of $[\mathbf{D}^H \mathbf{E}^{-1} \mathbf{D}]$. Note also that, in general, we will require the EVD of $\check{\mathbf{R}}_y = \mathbf{F}^{\frac{1}{2}} \mathbf{R}_y \mathbf{F}^{\frac{H}{2}} = \mathbf{F}^{\frac{1}{2}} \mathbf{D}^H \mathbf{R}_x \mathbf{D} \mathbf{F}^{\frac{H}{2}}$. Thus, in general, two N -dimensional EVDs will be required, one decomposing $\check{\mathbf{R}}_y$ and one decomposing $[\mathbf{D}^H \mathbf{E}^{-1} \mathbf{D}]$. When the element-space uncertainty set is a sphere, so that in (4), $\mathbf{E} = \frac{1}{\epsilon} \mathbf{I}$, then

$$\mathbf{F} = \frac{1}{\epsilon} (\mathbf{D}^H \mathbf{D})^{-1}. \quad (11)$$

In this case, if the DRT is orthogonal, \mathbf{F} in (11) reduces to $\mathbf{F} = \frac{1}{\epsilon} (\mathbf{D}^H \mathbf{D})^{-1} = \frac{1}{\epsilon} \mathbf{I}_N$. Thus, if the element-space set is a sphere and the DRT is orthogonal, then \mathbf{F} can be written analytically and only one EVD is required. Denoting $\hat{\mathbf{b}}_0$ as the solution to (7), we form the RDRCB weight vector as

$$\hat{\mathbf{w}}_{\text{RDRCB}} = \frac{\mathbf{R}_y^{-1} \hat{\mathbf{b}}_0}{\hat{\mathbf{b}}_0^H \mathbf{R}_y^{-1} \hat{\mathbf{b}}_0}. \quad (12)$$

The weight vector (12) operates on the reduced-dimension data. The weight vector that operates on the original element-space data is given by $\hat{\mathbf{w}}_{\text{RDRCB,ES}} = \mathbf{D} \hat{\mathbf{w}}_{\text{RDRCB}}$. An estimate of \mathbf{a}_0 can be formed as $\hat{\mathbf{a}}_0 = (\mathbf{D}^H)^{-1} \hat{\mathbf{b}}_0 = \mathbf{D} (\mathbf{D}^H \mathbf{D})^{-1} \hat{\mathbf{b}}_0$. Note that if $\hat{\mathbf{b}}_0 = \mathbf{D}^H \hat{\mathbf{a}}_0$, then $\hat{\mathbf{a}}_0 = \Pi_D \hat{\mathbf{a}}_0$, where Π_D is an orthogonal projection matrix onto the column space of \mathbf{D} . Given $\hat{\mathbf{a}}_0$, we form the RDRCB SOI power estimate as

$$\hat{\sigma}_{0,\text{RDRCB}}^2 = \frac{(\|\hat{\mathbf{a}}_0\|_2^2/M)}{\hat{\mathbf{b}}_0^H \mathbf{R}_y^{-1} \hat{\mathbf{b}}_0} = \frac{\hat{\mathbf{b}}_0^H (\mathbf{D}^H \mathbf{D})^{-1} \hat{\mathbf{b}}_0}{M \hat{\mathbf{b}}_0^H \mathbf{R}_y^{-1} \hat{\mathbf{b}}_0}. \quad (13)$$

3. DATA-DEPENDENT DIMENSION REDUCTION

Here, we consider Krylov methods that use the PoR and CG algorithms to compute the matrix that performs dimension reduction.

3.1. PoR (Non-Orthogonal) Krylov Basis

The standard PoR method for creating a Krylov DRT is to form

$$\mathbf{D} = \begin{bmatrix} \frac{\bar{\mathbf{a}}}{\|\bar{\mathbf{a}}\|_2} & \frac{\check{\mathbf{R}}_x \bar{\mathbf{a}}}{\|\check{\mathbf{R}}_x \bar{\mathbf{a}}\|_2} & \cdots & \frac{\check{\mathbf{R}}_x^{N-1} \bar{\mathbf{a}}}{\|\check{\mathbf{R}}_x^{N-1} \bar{\mathbf{a}}\|_2} \end{bmatrix}, \quad (14)$$

which can be formed iteratively. That is, starting with $\kappa_1 = \bar{\mathbf{a}}$, and $\mathbf{D}_1 = \frac{\bar{\mathbf{a}}}{\|\bar{\mathbf{a}}\|_2}$, for $i = 2, \dots, N$, calculate

$$\kappa_i = \hat{\mathbf{R}}_x \kappa_{i-1}, \quad (15)$$

$$\mathbf{d}_i = \frac{\kappa_i}{\|\kappa_i\|_2} \quad (16)$$

and

$$\mathbf{D}_i = [\mathbf{D}_{i-1} \quad \mathbf{d}_i] \quad (17)$$

. The cost of calculating κ_i from κ_{i-1} is $\mathcal{O}(M^2)$ and calculating \mathbf{d}_i is $\mathcal{O}(M)$. Thus, calculating the Krylov DRT costs $\mathcal{O}(NM[M+1])$ flops. The resulting Krylov DRT is non-orthogonal (NO) and therefore, to compute the NO-Krylov RDRCB, two N -dimensional EVDs will need computing, even if the original element-space set is spherical.

3.2. PoR Orthogonal Krylov Basis

In [9], the PoR orthogonal Krylov (O-Krylov) subspace technique was proposed and suggested for applications where the model order is highly variable and time-varying. To form the O-Krylov DRT, let $\kappa_1 = \bar{\mathbf{a}}$, $\mathbf{D}_1 = \bar{\mathbf{a}}/\|\bar{\mathbf{a}}\|_2$, and for $i = 2, \dots, N$, calculate

$$\kappa_i = \Pi_{\mathbf{D}_{i-1}}^\perp \hat{\mathbf{R}}_{\mathbf{x}} \kappa_{i-1} \quad (18)$$

$$\mathbf{d}_i = \frac{\kappa_i}{\|\kappa_i\|_2} \quad (19)$$

and

$$\mathbf{D}_i = [\mathbf{D}_{i-1} \quad \mathbf{d}_i] \quad (20)$$

, where $\Pi_{\mathbf{D}_{i-1}}^\perp = \mathbf{I} - \sum_{k=1}^{i-1} \mathbf{d}_k \mathbf{d}_k^H$ can be updated efficiently in $\mathcal{O}(M^2)$ operations using $\Pi_{\mathbf{D}_i}^\perp = \Pi_{\mathbf{D}_{i-1}}^\perp - \mathbf{d}_i \mathbf{d}_i^H$. Given $\Pi_{\mathbf{D}_{i-1}}^\perp$, updating κ_i and \mathbf{d}_i costs $\mathcal{O}(2M^2)$ and $\mathcal{O}(M)$. Thus, the calculation of one new column of \mathbf{D} costs $\mathcal{O}(3M^2 + M)$, so that calculation of the O-Krylov DRT costs $\mathcal{O}(NM[3M + 1])$, which is roughly three times more expensive than calculating the standard NO-Krylov DRT. Since the resulting DRT is orthogonal, as discussed earlier, for spherical uncertainty sets only one EVD is required to compute the RDRCB.

3.3. Conjugate Gradient Method

Using the approach outlined in [11], the CG DRT can be formed by setting, $\mathbf{d}_1 = \bar{\mathbf{a}}$, $\mathbf{r}_1 = -\bar{\mathbf{a}}$, and then for $i = 2, \dots, N$, update using

$$\alpha_i = -\frac{\mathbf{d}_i^H \mathbf{r}_i}{\mathbf{d}_i^H \hat{\mathbf{R}}_{\mathbf{x}} \mathbf{d}_i}, \quad (21)$$

$$\mathbf{r}_{i+1} = \mathbf{r}_i + \alpha_i \hat{\mathbf{R}}_{\mathbf{x}} \mathbf{d}_i, \quad (22)$$

$$\beta_i = \frac{\mathbf{d}_i^H \hat{\mathbf{R}}_{\mathbf{x}} \mathbf{r}_{i+1}}{\mathbf{d}_i^H \hat{\mathbf{R}}_{\mathbf{x}} \mathbf{d}_i} \quad (23)$$

and

$$\mathbf{d}_{i+1} = -\mathbf{r}_{i+1} + \beta_i \mathbf{d}_i. \quad (24)$$

The cost of computing $\hat{\mathbf{R}}_{\mathbf{x}} \mathbf{d}_i$ is $\mathcal{O}(M^2)$. Given $\hat{\mathbf{R}}_{\mathbf{x}} \mathbf{d}_i$, the cost of computing α_i is $\mathcal{O}(2M)$. Updating \mathbf{r}_{i+1} is $\mathcal{O}(M)$. The cost of computing β_i , given $\hat{\mathbf{R}}_{\mathbf{x}} \mathbf{d}_i$ and the denominator of α_i is $\mathcal{O}(M)$. Then, updating \mathbf{d}_{i+1} is $\mathcal{O}(M)$. Thus, the total cost to compute a new column of the CG DRT is $\mathcal{O}(M^2 + 5M)$. Thus, the total cost to calculate the CG DRT is $\mathcal{O}(NM[M + 5])$, which is almost the same as calculating the NO Krylov DRT. Since the CG DRT is NO, we would expect that we would need two EVDs to compute the CG-RDRCB. However, in the next section, we illustrate how a fast CG-based RDRCB can be obtained by exploiting that the CG DRT diagonalizes the SCM so that

$$\hat{\mathbf{R}}_{\mathbf{y}} = \mathbf{D}^H \hat{\mathbf{R}}_{\mathbf{x}} \mathbf{D} = \Lambda_{\text{CG}}, \quad (25)$$

where Λ_{CG} is a diagonal matrix and $\mathbf{D} = [\mathbf{d}_1 \dots \mathbf{d}_N]$ is the DRT matrix.

4. FAST CONJUGATE-GRADIENT RDRCB

Here, we illustrate how only one N -dimensional EVD is required to solve the CG-RDRCB under either spherical or non-degenerate uncertainty. In general, we will be solving

$$\min_{\mathbf{b}} \mathbf{b}^H \mathbf{R}_{\mathbf{y}}^{-1} \mathbf{b} \text{ s.t. } [\mathbf{b} - \bar{\mathbf{b}}]^H \mathbf{F} [\mathbf{b} - \bar{\mathbf{b}}] \leq 1. \quad (26)$$

Usually, at this stage one would transform with $\mathbf{F}^{\frac{1}{2}}$ to give a spherical uncertainty set. However, from (25), we observe that $\mathbf{R}_{\mathbf{y}}^{-1} = \Lambda_{\text{CG}}^{-1}$, so that (26) can be written as

$$\min_{\mathbf{b}} \mathbf{b}^H \Lambda_{\text{CG}}^{-1} \mathbf{b} \text{ s.t. } [\mathbf{b} - \bar{\mathbf{b}}]^H \mathbf{F} [\mathbf{b} - \bar{\mathbf{b}}] \leq 1. \quad (27)$$

Letting $\mathbf{M} = \Lambda_{\text{CG}}^{-\frac{1}{2}} \mathbf{D}^H \mathbf{E}^{-1} \mathbf{D} \Lambda_{\text{CG}}^{-\frac{H}{2}}$, $\check{\mathbf{b}} = \Lambda_{\text{CG}}^{-\frac{1}{2}} \mathbf{b}$ and $\check{\bar{\mathbf{b}}} = \Lambda_{\text{CG}}^{-\frac{1}{2}} \bar{\mathbf{b}}$, we can rewrite (27) as

$$\min_{\check{\mathbf{b}}} \check{\mathbf{b}}^H \check{\mathbf{b}} \text{ s.t. } [\check{\mathbf{b}} - \check{\bar{\mathbf{b}}}]^H \mathbf{M}^{-1} [\check{\mathbf{b}} - \check{\bar{\mathbf{b}}}] \leq 1. \quad (28)$$

We form the Lagrangian using the real Lagrange multiplier μ

$$L(\check{\mathbf{b}}, \mu) = \check{\mathbf{b}}^H \check{\mathbf{b}} + \mu \left([\check{\mathbf{b}} - \check{\bar{\mathbf{b}}}]^H \mathbf{M}^{-1} [\check{\mathbf{b}} - \check{\bar{\mathbf{b}}}] - 1 \right). \quad (29)$$

Setting $\frac{\partial L(\check{\mathbf{b}}, \mu)}{\partial \check{\mathbf{b}}^H} = \mathbf{0}$ yields

$$\hat{\mathbf{b}} = \left(\frac{\mathbf{M}}{\mu} + \mathbf{I} \right)^{-1} \check{\bar{\mathbf{b}}} = \check{\bar{\mathbf{b}}} - [\mu \mathbf{M}^{-1} + \mathbf{I}]^{-1} \check{\bar{\mathbf{b}}}, \quad (30)$$

where we have used the matrix inversion lemma to obtain the term after the second equality. Using (30) in the constraint equation in (28) yields

$$h(\hat{\mathbf{b}}, \mu) = \check{\bar{\mathbf{b}}}^H [\mu \mathbf{M}^{-1} + \mathbf{I}]^{-1} \mathbf{M}^{-1} [\mu \mathbf{M}^{-1} + \mathbf{I}]^{-1} \check{\bar{\mathbf{b}}}^H. \quad (31)$$

Letting $\mathbf{M} = \mathbf{U} \Lambda \mathbf{U}^H$ denote the EVD of \mathbf{M} , where $\Lambda = \text{diag} \{ \lambda_1 \dots \lambda_N \}$ is a diagonal matrix containing the eigenvalues in non-increasing order on its main diagonal and \mathbf{U} contains the associated eigenvectors, we can write (31) as

$$h(\hat{\mathbf{b}}, \mu) = \sum_{n=1}^N \frac{\lambda_n |c_n|^2}{(\mu + \lambda_n)^2}, \quad (32)$$

where c_n is the n th element of $\mathbf{c} = \mathbf{U}^H \check{\bar{\mathbf{b}}}$. Since we can write $\mathbf{M} = \mathbf{M}^{\frac{1}{2}} \mathbf{M}^{\frac{H}{2}}$, where $\mathbf{M}^{\frac{1}{2}} = \Lambda_{\text{CG}}^{-\frac{1}{2}} \mathbf{D}^H \mathbf{E}^{-\frac{1}{2}}$, we know that \mathbf{M} is non-negative definite [15, 16] and therefore, it has non-negative eigenvalues. Thus, $h(\hat{\mathbf{b}}, \mu)$ is a monotonically decreasing function of $\mu > 0$. For $\mu = 0$, we obtain

$$h(\hat{\mathbf{b}}, 0) = \check{\bar{\mathbf{b}}}^H \mathbf{M}^{-1} \check{\bar{\mathbf{b}}} = \bar{\mathbf{b}}^H [\mathbf{D}^H \mathbf{E}^{-1} \mathbf{D}]^{-1} \bar{\mathbf{b}} = \bar{\mathbf{b}}^H \mathbf{F} \bar{\mathbf{b}}. \quad (33)$$

Note that, to exclude a non-trivial solution we require that, $\bar{\mathbf{b}}^H \mathbf{F} \bar{\mathbf{b}} > 1$. Since we require $h(\hat{\mathbf{b}}, \mu) = 1$, it is clear that $\mu \neq 0$. Further, it is clear that $\lim_{\mu \rightarrow \infty} h(\hat{\mathbf{b}}, \mu) = 0$, therefore, there is a unique solution $\mu > 0$ to $h(\hat{\mathbf{b}}, \mu) = 1$, which can be found, e.g., by Newton search.

Once μ has been found, $\hat{\mathbf{b}}$ is found using (30) and the solution to (27) is formed as $\hat{\mathbf{b}}_0 = \Lambda_{\text{CG}}^{\frac{1}{2}} \hat{\mathbf{b}}$. We can use $\hat{\mathbf{b}}_0$ and $\mathbf{R}_{\mathbf{y}}^{-1} = \Lambda_{\text{CG}}^{-1}$ in (12) to form the adaptive weights. To form the power estimate using (13), we need to evaluate $[\mathbf{D}^H \mathbf{D}]^{-1}$. If the uncertainty set is spherical, then we can evaluate this quantity from the EVD of \mathbf{M} and Λ_{CG} , which are already available. For a general, non-degenerate ellipsoid this quantity will need computing.

Fig. 1 shows the relative complexities as N is increased from 1 to M , for $M = 320$, illustrating that the CG-based algorithms are significantly cheaper than the other methods.

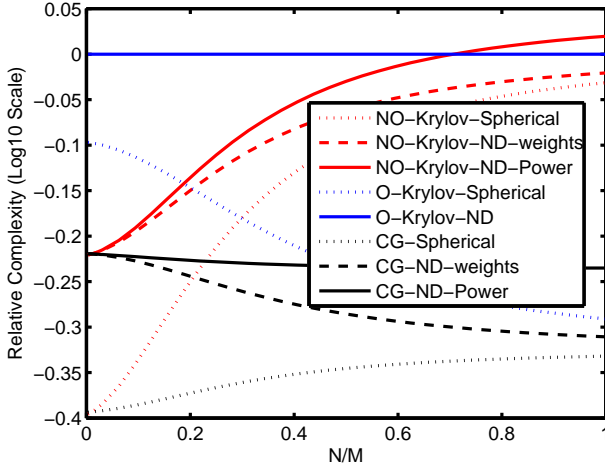


Fig. 1. Relative complexities of different data-dependent RDRCBs.

5. NUMERICAL EXAMPLES

In this section, we assess the performance of the proposed algorithms through numerical examples. For an $M = 320$, $\lambda/2$ -spaced planar array with $M_h = 40$ elements in a row and $M_v = 8$ rows, we simulated data with covariance matrix $\mathbf{R}_x = \sigma_0^2 \mathbf{a}_0 \mathbf{a}_0^H + \mathbf{Q}_x$, with $\mathbf{Q}_x = \sum_{i=1}^d \sigma_i^2 \mathbf{a}_i \mathbf{a}_i^H + \sigma_s^2 \mathbf{I} + \sigma_{\text{iso}}^2 \mathbf{Q}_{\text{iso}}$, where \mathbf{Q}_x consists of terms due to d zero-mean uncorrelated interferences, where for the i th interferer σ_i^2 and \mathbf{a}_i denote the source power and ASV, a term modeling sensor noise $\sigma_s^2 \mathbf{I}$, with sensor noise power σ_s^2 , and a term modeling an isotropic ambient noise $\sigma_{\text{iso}}^2 \mathbf{Q}_{\text{iso}}$, with power σ_{iso}^2 . The isotropic noise covariance is given by $[\mathbf{Q}_{\text{iso}}]_{m,n} = \text{sinc}[\pi g_{mn}]$, where g_{mn} is the distance between the m th and n th sensors in units of wavelength. The i th source (SOI or interference) ASV is simulated according to $\mathbf{a}_i = \mathbf{a}(\bar{\theta}_i + \delta_i) + \sigma_{e,i} \mathbf{e}_i$, where \mathbf{e}_i is a zero-mean complex circularly symmetric random vector with unit norm. When $\delta_i \neq 0$ an AOA error exists and when $\sigma_{e,i} \neq 0$, an arbitrary error exists. We assume azimuth and elevation beams spaced at $1/M_h$ and $1/M_v$ in cosine space and, using the methods described in [17], design tight-spherical uncertainty sets and non-degenerate minimum volume ellipsoidal (NDMVE) sets based on the expected AOA errors given the spacing of the beams. Fig. 2 shows SINR versus SNR for....??? The results show that the CG and O-Krylov results are the same, whilst the NO-Krylov results diverge for very high SNRs. It is clear that the robust RDRCB version exploiting spherical or non-degenerate NDMVE sets, provide much better robustness at high SNRs compared to the standard MVDR-based implementations.

6. REFERENCES

- [1] S. D. Somasundaram, "Reduced Dimension Robust Capon Beamforming for Large Aperture Passive Sonar Arrays," *IET Radar, Sonar Navig.*, vol. 5, no. 7, pp. 707–715, Aug. 2011.
- [2] S. A. Vorobyov, A. B. Gershman, and Z.-Q. Luo, "Robust Adaptive Beamforming Using Worst-Case Performance Optimization: A Solution to the Signal Mismatch Problem," *IEEE Trans. Signal Process.*, vol. 51, no. 2, pp. 313–324, Feb. 2003.
- [3] P. Stoica, Z. Wang, and J. Li, "Robust Capon Beamforming," *IEEE Sig. Process. Lett.*, vol. 10, no. 6, pp. 172–175, Jun. 2003.
- [4] J. Li, P. Stoica, and Z. Wang, "On Robust Capon Beamforming

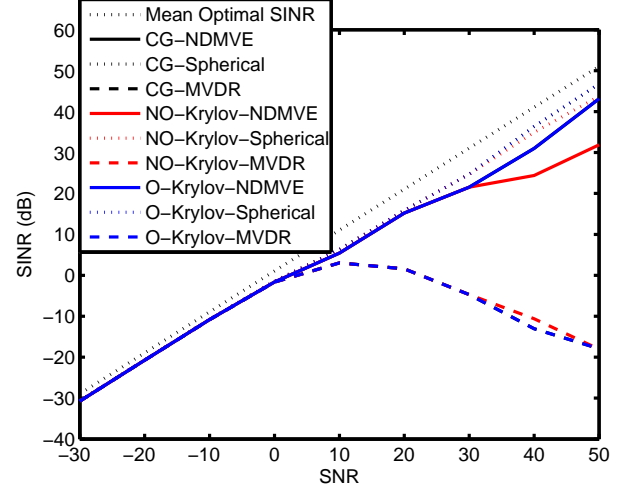


Fig. 2. SINR versus SNR for the algorithms analyzed.

and Diagonal Loading," *IEEE Trans. Signal Process.*, vol. 51, no. 7, pp. 1702–1715, Jul. 2003.

- [5] R. G. Lorenz and S. P. Boyd, "Robust Minimum Variance Beamforming," *IEEE Trans. Signal Process.*, vol. 53, no. 5, pp. 1684–1696, May 2005.
- [6] J. Li and P. Stoica, *Robust Adaptive Beamforming*, Wiley, New York, 2005.
- [7] S. D. Somasundaram, "A Framework for Reduced Dimension Robust Capon Beamforming," in *Proc. IEEE Workshop on Statistical Signal Processing*, Nice, France, Jun. 28–30 2011, pp. 425–428.
- [8] M. L. Honig and J. S. Goldstein, "Adaptive reduced-rank interference suppression based on the multistage wiener filter," *IEEE Trans. on Communications*, vol. 50, pp. 986–994, 2002.
- [9] H. Ge, I. P. Kirsteins, and L. L. Scharf, "Data Dimension Reduction Using Krylov Subspaces: Making Adaptive Beamformers Robust to Model Order-Determination," in *31st IEEE International Conference on Acoustics, Speech and Signal Processing*, Toulouse, Fr, May. 2006, vol. 4, pp. 1001–1004.
- [10] M. Haardt R. C. de Lamare and R. Sampaio-Neto, "Blind adaptive constrained reduced-rank parameter estimation based on constant modulus design for CDMA interference suppression," *IEEE Trans. Sig. Proc.*, vol. 56, no. 2, pp. 2470–2482, 2008.
- [11] G. Dietl, "Conjugate Gradient Implementation of Multi-Stage Nested Wiener Filter for Reduced-Dimension Processing," in *MSc. Dissertation, Munich University of Technology*, Munich, Germany, May 2001.
- [12] L. Wang and R. C. de Lamare, "Constrained adaptive filtering algorithms based on conjugate gradient techniques for beamforming," *IET Signal Processing*, vol. 4, pp. 686–697, 2010.
- [13] R. C. de Lamare and R. Sampaio-Neto, "Adaptive reduced-rank mmse filtering with interpolated fir filters and adaptive interpolators," *IEEE Sig. Proc. Letters*, vol. 12, pp. 177–180, 2005.
- [14] R. C. de Lamare and R. Sampaio-Neto, "Adaptive reduced-rank processing based on joint and iterative interpolation, decimation, and filtering," *IEEE Trans. Sig. Proc.*, vol. 57, pp. 2503–2514, 2009.

- [15] G. Strang, *Linear Algebra and its Applications (3rd Edition)*, Harcourt Brace Jovanovich, 1988.
- [16] P. Stoica and R. L. Moses, *Spectral Analysis of Signals*, Prentice Hall, 2005.
- [17] S. D. Somasundaram, A. Jakobsson, and N. H. Parsons, "Robust and Automatic Data-Adaptive Beamforming for Multi-Dimensional Arrays," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 11, pp. 4642–4656, Nov. 2012.